

Title (en)

DEEP LEARNING MODEL TRAINING METHOD AND SYSTEM

Title (de)

VERFAHREN UND SYSTEM ZUM TRAINIEREN VON TIEFENLERNMODELLEN

Title (fr)

PROCÉDÉ ET SYSTÈME D'ENTRAÎNEMENT DE MODÈLE D'APPRENTISSAGE PROFOND

Publication

EP 3889846 A4 20220601 (EN)

Application

EP 19910606 A 20190124

Priority

- CN 201910041235 A 20190116
- CN 2019072895 W 20190124

Abstract (en)

[origin: EP3889846A1] This application provides a deep learning model training method. The method includes: generating N first gradient sets in BP calculation in the jth iteration of N deep learning models; adjusting a communication sequence of gradients included in each of the first gradient sets, where the gradients included in each of the first gradient sets are not sent to parameter storage space in a generation sequence of the gradients included in each of the first gradient sets; and separately sending, according to an adjusted communication sequence of the gradients, the gradients included in each of the N first gradient sets to the parameter storage space. According to the method, a sequence of transmitting a gradient obtained in a current iteration process to the parameter storage space is adjusted, to increase training efficiency of the deep learning model.

IPC 8 full level

G06N 3/08 (2006.01); **G06N 3/04** (2006.01); **G06N 3/063** (2006.01)

CPC (source: EP US)

G06N 3/045 (2023.01 - EP US); **G06N 3/063** (2013.01 - EP); **G06N 3/084** (2013.01 - EP US)

Citation (search report)

- [X] PAN XINGHAO ET AL: "Revisiting Distributed Synchronous SGD", 21 March 2017 (2017-03-21), XP055781228, Retrieved from the Internet <URL:https://arxiv.org/pdf/1604.00981.pdf> [retrieved on 20210302]
- [A] GENG TONG ET AL: "A Framework for Acceleration of CNN Training on Deeply-Pipelined FPGA Clusters with Work and Weight Load Balancing", 2018 28TH INTERNATIONAL CONFERENCE ON FIELD PROGRAMMABLE LOGIC AND APPLICATIONS (FPL), IEEE, 27 August 2018 (2018-08-27), pages 394 - 3944, XP033462513, DOI: 10.1109/FPL.2018.00074
- [A] TAL BEN-NUN ET AL: "Demystifying Parallel and Distributed Deep Learning: An In-Depth Concurrency Analysis", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 26 February 2018 (2018-02-26), XP081074598
- See also references of WO 2020147142A1

Designated contracting state (EPC)

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

DOCDB simple family (publication)

EP 3889846 A1 20211006; **EP 3889846 A4 20220601**; CN 111788585 A 20201016; CN 111788585 B 20240412; US 2021342696 A1 20211104; WO 2020147142 A1 20200723

DOCDB simple family (application)

EP 19910606 A 20190124; CN 2019072895 W 20190124; CN 201980000128 A 20190124; US 202117376722 A 20210715